DOCUMENT RESUME

ED 376 181 TM 022 169

AUTHOR	Smith, Mary Lee; And Others
TITLE	What Happens When the Test Mandate Changes? Results of a Multiple Case Study. Project 3.2 State Accountability Models in Action.
INSTITUTION	National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Mar 94
CONTRACT	R117G10027
NOTE	75p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-8, 1994).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Accountability; Case Studies; Criterion Referenced Tests; Data Collection; *Educational Assessment; Educational Change; Elementary Secondary Education; *Program Implementation; State Legislation; *State Programs; *Testing Programs; *Test Use
IDENTIFIERS	Arizona; *Arizona Student Assessment Program; Iowa Tests of Basic Skills; *Performance Based Evaluation

ABSTRACT

The 1992-93 school year marked the first year of implementation of the statewide mandate of the Arizona Student Assessment Program (ASAP), which revised previous legislation. Iowa Test of Basic Skills testing requirements were restricted to grades 2 and 7, and districts were allowed greater flexibility in their own testing, which had previously been almost exclusively through criterion-referenced tests. Performance assessments were supported by the new legislation, and teachers generally saw the ASAP as a low-stakes assessment in line with educational trends. The implementation of the ASAP and its changes were studied in four elementary schools during the first implementation year in a multiple case-study design with various data collection methods. Results indicate that local responses to the ASAP were varied and that differences in implementation were sighificant. Common among the sites was the belief that testing from an outside agency is still separate from instruction and is an add-on to normal school operation. An appendix contains a cross-site data matrix. (Contains 14 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 376 181

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

National Center for Research on
Evaluation, Standards, and Student Testing

Final Deliverable – March 1994

Project 3.2 State Accountability Models in Action

What Happens When the Test Mandate Changes? Results of a Multiple Case Study

Mary Lee Smith, Project Study Director

U.S. Department of Education
Office of Educational Research and Improvement
Grant No. R117G10027 CFDA Catalog No. 84.117G

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Graduate School of Education
University of California, Los Angeles
Los Angeles, CA 90024-1522
(310) 206-1532

TM022169

BEST COPY AVAILABLE

The work reported herein was supported under the Educational Research and Development Center Program, cooperative agreement number R117G10027 and CFDA catalog number 84.117G, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the position or policies of the Office of Educational Research and Improvement or the U.S. Department of Education.

WHAT HAPPENS WHEN THE TEST MANDATE CHANGES? RESULTS OF A MULTIPLE CASE STUDY[1]

Mary Lee Smith, Audrey J. Noble, Marilyn Cabay,

Walt Heinecke, M. Susan Junker, and Yvonne Saffron

CRESST/Arizona State University

Introduction

The academic year 1992-1993 marked the first year of implementation of the statewide mandate known as the Arizona Student Assessment Program (ASAP), which was authorized by the Arizona Revised Statutes 15-741-744 of 1990. This bill revised previous legislation, which had mandated testing every pupil every spring on the Iowa Tests of Basic Skills (ITBS) in Grades 2-8 and the Test of Academic Performance (TAP) in Grades 9-11. The mandate also included the requirement that districts develop and administer tests to determine if schools were meeting the Arizona Essential Skills, the statewide curriculum framework. ASAP reduced the ITBS testing requirements to Grades 2 and 7 (and TAP to 11) and moved the testing date to the fall. District testing, which heretofore had been almost exclusively by criterion-referenced methods, was allowed greater flexibility: Districts could continue with CRTs, use portfolio assessments, or administer and locally score the new performance assessments using Forms A, B, and C. Form D was designed to be the on-demand or audit form of the performance assessment. It was administered during March to pupils in Grades 3, 8 and 12, with standardized administration rules and procedures. Rubrics for scoring the performance test were used, at central scoring sites, by teachers trained by state officials

[1] This work was also reported in a paper presented at the annual meeting of the American Educational Research Association, New Orleans, Louisiana, April 7, 1994.

and representatives of the test developers, Riverside Press and Measurement, Inc. Scores were reported by student to schools and districts, and by school and district to general audiences. District average scores were one part of a state-required Report Card, also including ITBS and district test scores, all referenced to the Arizona Essential Skills. That is, each district had to submit a District Assessment Plan (DAP), specifying a mastery level on each of the Essential Skills and reporting the percentage of pupils who had attained that level, as indicated by the collection of assessment results. Although ASAP included all these components (performance assessment Form D, ITBS/TAP, district testing, DAP, School Report Cards), most people used the term "ASAP" to refer only to the performance assessment itself.

Like any mandate, ASAP was designed to solve what policy makers perceive to be a problem. The perceived solution to the problem lies in requiring some uniform action on the part of its agents (McDonnell & Elmore, 1987). In the case of ASAP, at least two categories of problem were in the minds of policy makers. In Noble (1994) and Noble and Smith (in press), we reported results of a policy study in which the images and beliefs of policy makers and state officials instrumental in the ASAP mandate were examined. Some of these individuals conceived of ASAP as a means of improving Arizona schools by moving them toward a more ambitious and integrated form of curriculum and pedagogy; that is, toward holistic teaching and higher order thinking or cognitive-constructivist learning. Others, however, conceived of ASAP as a means of making schools more accountable for achievement results, specifically focusing the school's attention more intensively on the Arizona Essential Skills.

According to officials of the Arizona Education Association and data collected early in 1992-1993, most teachers considered ASAP to be benign, supportive of educational trends, and a low-stakes assessment. Just as teachers thought of "ASAP" as equated with the performance assessment, they also interpreted the mandate as improving instruction toward holism and cognitive-constructivism. This interpretation was supported by initial information teachers received in state and regional conferences and training sessions run by the Arizona Department of Education (ADE). A pilot administration, conducted during academic year 1991-1992, reinforced this view. Most teachers who examined the pilot test material or participated in

the pilot administration seemed to think that ASAP was "a step in the right direction." By this, they meant that the performance test was a substantial improvement over the ITBS and supported a variety of instructional practice that they appreciated. As the findings in this report will show, this view was repeatedly challenged over the 1992-1993 year. At the end of the year, the ADE published district test results, and the newspaper distributed them, adding editorial comments about the failings of public schools. The ADE administrators used ASAP results for the same purpose, which altered many of the teachers' views about the function of ASAP.

Conceptual Context of the Study

The study reported herein is part of a larger project, "What Happens When the Test Mandate Changes?" The project encompasses three years of data collection on the consequences in Arizona of the implementation of ASAP. Several levels of analysis are covered in the project as a whole. The policy study (Noble, 1994) analyzes the images, beliefs, and values of policy makers and administrators as they reflect on the policy change, its antecedents and consequences. The present study addresses the consequences of the change in mandate in four Arizona elementary schools during the first year of implementation. During academic year 1993-1994 (the second year of policy implementation), we are extending the findings and testing the models through focus group interviews and survey methods. The report of the project as a whole will focus on the interplay of policy and practice over two years of policy implementation and local reactions.

Focusing on the interplay of policy and practice is a decision that comes from our conceptual framework. We drew on Rein (1983) and Weatherly and Lipsky (1978) for ideas about where to look for evidence about the effects of school policy making. From reading these works, we were committed to the idea that definitions of the situation (the images of the problems a given policy should solve as well as the characteristics of pupils, teachers, curriculum, assessment, and educational change) held by policy makers and shapers are translated imperfectly by practitioners. Teachers and principals redefine and reinterpret the messages about policy that they receive. They then act—adapt, teach, learn, evaluate—according to their own definitions of the situation

(Blumer, 1986). This study, therefore, is symbolic interactionist in conceptual framework and interpretivist in research methodology.

Specifically, we draw on Erickson (1986), as well as Miles and Huberman (1984) for our research methods. That is, to understand action and practice, we believe that the researcher must engage directly in the local scene, spend sufficient time to understand action in its specific social context and gain access to participant meanings, and show how these meanings-in-action evolve over time. Without careful grounding in local cases, a more general understanding is impossible.

This study also draws on previous research on the role of mandated testing. An earlier qualitative study (Smith, Edelsky, Draper, Rottenberg, & Cherland, 1990) showed that the previous test mandate in Arizona, which involved the high-stakes use of the ITBS, had effects such as narrowing curriculum, promoting test-like instructional methods, reducing time for ordinary instruction, deskilling and demoralizing teachers, and leading to inappropriate test preparation practices. A review of related research (Smith, 1993) showed that similar effects have been experienced in other states and settings having high-stakes accountability programs. The question unanswered by extant research is whether assessments that differ in form from the traditional, norm- or criterion-referenced standardized tests would produce similar reactions and effects.

Proponents of performance assessment believe that what is assessed is what gets taught. Therefore, the argument goes, mandating an assessment that requires integrated curriculum (e.g., reading and math) and higher order thinking and problem solving on the part of pupils will drive schools and teachers to align their offerings so that pupils will be able to perform adequately (cf. Resnick, 1989). This is the essence of measurement-driven reform: that building a better test will drive schools toward more ambitious goals and reform them toward a curriculum and pedagogy geared more toward thinking and less toward rote memory and isolated skills—the shift from behaviorism to cognitive-constructivism. The present study represents an attempt to understand what happens during the initial year of implementation of such an assessment, which state officials have termed "the best we know about assessment and pupil learning."

Methods of the Study

The research design chosen to address this issue is the multiple-case-study design (Miles & Huberman, 1984). This design is based on the rationale that understanding complex organizations such as schools requires long-term and close-up examination of local practice within bounded social settings. The actions of participants faced with a new government mandate can only be understood in the specific context in which they occur and referenced to the meanings held by those participants. The researcher aiming to understand these meanings must have access, over an extended period of time, to the classrooms and offices in which participants' definitions of the situation (mandated assessments, in this case) evolve and get worked out in actions. Do they actually provide the type of instruction geared to the ASAP performance test? Do they have the knowledge they need to adapt, or do they have the intention to do so? What is the meaning of the ASAP to teachers and others in schools? Getting evidence to answer questions such as these requires more than snapshot observations and prespecified questionnaire items. Thus, the qualitative case study is the best design. The decision to do more than one case study was not made because four is closer to the population of schools than one. Nor is there any intent to evaluate the four schools comparatively. The rationale for drawing multiple cases is that one case provides interpretive context for the others. A case study researcher typically immerses herself in a single site and tries to understand everything there is to know about it. Holistic understanding, however, sometimes produces the holistic fallacy. Things unobserved in that setting are often not considered as salient; observed phenomena and events may be mistakenly seen as causal. Seeing two case studies in parallel can alert the two researchers of features taken for granted or overlooked in one. In the present study, for example, the influence of the district's philosophical support of ASAP was overlooked by the researcher in her within-case analysis. Simply because it was taken for granted by everyone in the site, she failed to observe the potential influence of this condition. Yet when her case was held up against another site, in which the district administration was not supportive of the mandate, the importance of the factor in explaining the relative success of the mandate in the two sites became obvious.

Four cases were chosen for the study. The number was determined by the resources available to support four graduate students for the year. Only elementary schools were chosen, because of the need to contrast the effects of the new mandate with the previous one studied by Smith et al. (1990). The decision of which sites to select was made based on the desirability of varying cases across economic and social resources and prior history of testing demand (the importance of test results historically in the district). Thus we tried to find schools with greater and lesser economic resources, serving advantaged and disadvantaged students, and located in urban, rural, and suburban settings. In addition, we made use of contacts and acquaintances that would help us to access particular schools and districts.

All schools we contacted and requested permission to study responded positively. The four sites where we conducted case studies were (a) Valor, a rural school with a low resource base, serving mostly poor and minority pupils in a K-8 district; (b) Franklin, an urban school with a relatively high resource base, serving mostly poor and minority pupils in a K-8 district; (c) Pines, a suburban school with an ethnically and economically diverse student body, in a large, K-8, resource-advantaged district with high test demand characteristics; and (d) Hilldale, a suburban school serving mostly Anglo and advantaged pupils, in a large, K-12, resource-advantaged district with moderate test demand characteristics. Additional information on the descriptive characteristics of the four sites is available in the case studies themselves and summarized in the Cross-Site Data Matrix (see Appendix). All names used in the study are pseudonyms. District and school personnel were promised confidentiality.

Five researchers were selected to conduct the case studies. Audrey Noble, assigned to Valor, is a fourth-year graduate student in the doctoral program in educational leadership and policy studies. In addition to her case study, she acted as research coordinator for the others. Suzii Junker, a third-year student in the doctoral program in reading, conducted the study at Hilldale. Walt Heinecke, a third-year student in the doctoral program in educational leadership and policy studies, studied Pines. Marilyn Cabay and Yvonne Saffron collaborated on the study at Franklin. Cabay and Saffron are fourth-year students in the doctoral program in school psychology. All five of the researchers had at least two courses in qualitative research at the time of the

study and had produced independent studies as part of their degree programs. All are highly experienced in various educational roles: classroom teacher, counselor, school administrator, school psychologist, testing coordinator. All five brought unique perspectives to their research role; yet consistency across researcher perspectives was maintained in several ways. First, a common design for data collection and common definitions of researcher roles were shared. Second, the theoretical framework focused researchers' attention on common aspects of the sites (the images held by the participants of pupil, teacher, learning, curriculum, assessment, and school structure). Third, monthly meetings of the researchers were held to address issues raised and problems at the separate sites, share memos and working papers, and the like. Fourth, the work of the researchers was supervised by Mary Lee Smith, who monitored the adequacy of data collection and analysis procedures. Finally, drafts of the four case studies were read by all members of the research team, and reactions were incorporated into the case studies by the researchers to add to the overall fit of the cases together and provide the interpretive context of each case to the others.

Data Collection

Each case study involved the following data collection methods. The unit of study was defined as the classroom within the school. The four participating schools provided the researchers with access to faculty meetings and other school events, direct observation of one third-grade and one fourth-grade class (except for Hilldale Elementary, in which a combined third/fourth-grade class was the primary participant), interviews with third-grade and fourth-grade teachers, and documents relevant to ASAP, curricula, and local testing programs. This access extended through the academic year 1992-93. Informal contact between researchers and participating teachers was maintained through 1993. The choice of third- and fourth-grade classes was based on the state mandate of ITBS testing in fourth grade during the month of October and ASAP performance testing in third grade in March. The design of observations followed from this schedule, with observation occasions clustered in the fourth-grade classes in the fall and the third-grade classes in the spring. The working design called for researchers to be in the targeted classrooms one day each week normally and twice per week immediately before, during, and after the testing events. They deviated from the schedule

when necessary to capture activities relevant to the research questions in the rest of the school or district. For example, the researcher at Hilldale accompanied the teacher whose class she usually observed when the teacher attended a training session on scoring of the performance test. The researcher at Valor branched out to classes other than the one chosen in the design so that she could understand the relative authority of teachers, principal, and district officials in determining curriculum choices.

The researchers played the role of "more observer-than-participant" (Gold, 1958), developing cordial, nonevaluative, and trusting relationships with the teachers and school staff. No problems with access were experienced at the schools over the year's data collection. However, project policy about confidentiality and ownership of the data had to be clarified and reiterated with officials in one of the districts. Our position was to maintain confidentiality and protection of the identity and perspectives of the participants with whom we dealt most directly—the teachers and principals. District officials would have access to only those data either that shielded the identity of the participants or that the participants had cleared for publication.

Observation occasions of school and classroom activities were aimed at understanding the role of testing in context, the meaning of mandated testing to teachers and school staff, test preparation for mandated tests, and the relationship of mandated testing to curriculum, pedagogy, and school structures. The conceptual framework of the study provided the focus for observations. That is, the researchers kept in mind the need to attend to, besides the normal, everyday life of the classrooms, incidents that shed light on the images held by participants of pupil, teacher, learning, assessment, and school structure. Researchers kept detailed notes of what they observed, transcribed their working notes, and submitted the write-ups in text files to the research coordinator. These were reviewed periodically to make sure the researchers were preserving the necessary level of concrete detail and recording material relevant to the research questions and conceptual framework. Monthly meetings of the researchers were held to coordinate insights and keep everyone on target.

By design, the researchers conducted formal interviews with the principal and teachers whom they observed and focus group interviews with remaining third- and fourth-grade teachers in the school. In addition,

interviews with district officials were conducted to understand the district perspectives on assessment and the organizational climate of the districts. The interview agenda and key questions and probes were developed by the research director and coordinator to generate data according to the conceptual framework. For example, teachers were asked questions such as: "The state believes that the new testing program will promote a new kind of instruction. Other than knowing what the test covers and how to administer it, what are the things a teacher needs to know to teach in the manner that ASAP promotes?" Because these interviews fit a qualitative approach to research, the exact wording and sequence of questions varied. It was more important to elicit the meanings the assessment had for participants than to standardize questions. The interviewees were encouraged to tell their own stories in their own words, the researchers using those words to construct probes so that the agenda could be addressed. For example, the probe for the question stated above might attempt to elicit information on the kinds and amounts of professional development the teachers had already experienced or believed to be important precursors of ASAP-related instruction. The agenda was drawn from the conceptual framework and emerging issues in the study as a whole. Interviews were tape-recorded and the tapes transcribed.

Researchers at the four sites also collected documents and artifacts. For example, some teachers voluntarily provided work samples from students in ASAP-related activities and journals in which students described their reaction to assessments. Curriculum guides, text samples, work sheets and instructional packets, detailed samples of district tests and test results, information sent to parents, notices of meetings and training sessions, and the like also supplemented the observation and interview data.

Within-Site Data Analysis

The researchers coded their data according to the categories in the project conceptual framework as well as categories emerging from their site. For example, every instance of data that plausibly referred to or illustrated a teacher's image of the curriculum was so coded for subsequent retrieval. Or, a district administrator's contention that district CRTs were a more appropriate standard for achievement than ASAP results would have been coded as "image of testing." In addition, local issues, such as the conflict among third- and

fourth-grade teachers at Franklin about the value of moving to ASAP-like instruction, produced the inductively-derived category "Grade-level isolation/ conflict." Researchers were encouraged to use qualitative analysis computer programs, such as Ethnograph and Hyperqual, to identify, mark, index, and retrieve data that instantiated the categories. They wrote memos periodically to define the categories and document their thinking processes as they analyzed their data. Finally, they wrote assertions and produced vignettes to support the assertions. According to Erickson (1986), assertions are statements that researchers inductively derive by reading and re-reading the record and data. These statements are inferences about the meaning of the evidence. For example, one of the assertions from the study of the Valor site follows: "Although performance assessment is meant to encourage the social nature of learning, learned attitudes and behaviors (prior knowledge) regarding testing persist. Teachers and students respond to the function of assessment rather than the form. Testing for teachers and students remains a solitary, inactive, and structured experience." Vignettes had two functions: to describe a particular slice of life in the setting and to illustrate the basis in data from which the assertion was derived (Erickson, 1986). Thus, the vignette that accompanied the above-quoted assertion vividly describes how teachers prepared for and administered both ASAP and ITBS. The style and tone of ASAP administration resembled that of ITBS but contrasted with that of regular instruction.

Researchers established the warrant for their assertions by looking closely for disconfirming instances, and checking that the assertions had sufficient confirming data of varying methods (e.g., observations vs. interviews). In addition, drafts of the assertions and vignettes of each case study were read by the other researchers, the coordinator and director. Revisions were made based on this feedback. Then, the researchers completed the case studies (Smith et al., 1994), providing their overall perspective about the role of mandated testing in their respective sites.

Cross-Site Data Analysis

The existence and use of the conceptual framework for the study as a whole, the monthly meetings, and supervision of researchers increased the likelihood that the separate case studies would have enough elements in

common to enable cross-site analysis. The final meeting of the research team to discuss the case studies was tape-recorded to preserve a record of the ideas generated. This meeting served two analytic purposes. First, each case was used as interpretive context for the others. That is, elements that had been overlooked in one site became highlighted by comparing cases. For example, at Hilldale, district testing was simply not an issue, and the researcher at that site had consequently ignored it. At Franklin and Pines, however, the district testing program has profound impact on what happens to ASAP-relevant instruction. Through this comparison, a hole in the Hilldale account was readily identified and rectified. Second, it treated the researchers as informants in the sense that, after a year of data collection, they "knew" much more about the educational and social context than they could have possibly included in the case study. The director and coordinator could then ask them to summarize information on issues of cross-site interest. For example, a quick reading of data and a few phone calls produced data on the missing element from the Hilldale account on the role of district testing.

The analysis of qualitative data is fundamentally a process of thinking and progressive problem solving (Erickson, 1986), with only a crude set of tools and procedures. The conceptual framework yielded categories such as Image of the Pupil. Data had been gathered that allowed us to generate assertions within each site about the Image of the Pupil that seemed to be held by teachers and district officials. In addition, we had evolved a set of working hypotheses, or plausible accounts and explanations, for how the change in mandated testing was working out at each site, that is, what particular barriers and facilitating conditions seemed to be responsible for local reactions. Furthermore, we understood that audiences for this report would be interested in the formal characteristics of each site (e.g., the degree of pupil disadvantage) and would need a variety of information to make their own interpretations of the data. From these considerations, we developed a set of dimensions for the cross-site matrix. Our aim was to provide data in the matrix that would reduce the sheer quantity of information to a manageable level without resorting to high-level abstractions or losing the sense of grounding and authenticity that case studies can provide.

Based on the above considerations, the Cross-Site Data Matrix was constructed. The elements in each cell are short summaries, paraphrases, or

characterizations of the particular site on the selected dimensions. These characterizations were constructed by the research director and submitted to the case study researchers for their substantive and editorial comments.

The Cross-Site Data Matrix is placed in the Appendix. The dimensions of the matrix are as follows:

DISTRICT CHARACTERISTICS
- Size and Organization (K-8)
- Resource Base
- Organizational Culture
- District Testing Model
- Test Demand (Hi/Low Stakes)
- Knowledge/Commitment of Officials to ASAP-like Instructional Principles
- Belief in the Permanence of ASAP
- Image of Pupil
- Image of Teacher
- Image of Assessment
- Image of Curriculum
- Reaction to ASAP Results
- Prospects for Second-Year Changes

PUPIL CHARACTERISTICS
- SES
- Language Dominance
- Ethnic Composition

COMMUNITY CHARACTERISTICS
- Size and Type of Community
- SES
- Parent Participation and Interest in Scores

SCHOOL CHARACTERISTICS
- School Organization and Size
- School Structure
- Grade-level Isolation/Conflict
- Role of Principal
- Principal Accommodation/Resistance
- Curriculum/Texts
- Test Burden
- Test Preparation
- Presence of ASAP Key Gatekeeper

TEACHER CHARACTERISTICS (Focal Teachers)

- Experience
- Commitment to Holistic, Thinking Instruction
- Prior Knowledge of ASAP-like Instruction
- Opportunities for Relevant Professional Development
- Familiarity With Performance Test, Rubrics, Essential Skills
- The Professional Life of Teachers
- Image of Pupil
- Image of Teacher
- Image of Curriculum
- Image of Assessment
- Test Preparation (activities engaged in)
- Time to Reflect, Experiment, Collaborate
- Perception of/Reaction to Test Stakes
- Accommodation/Resistance

The process of arraying data in the Cross-Site Data Matrix stimulated further thinking about what elements were most salient in accounting for the differences among the cases in response to the mandate. In constructing the Analytic Matrix (Figure 1), we started with a working assumption (analyzed and critiqued in Noble, 1994) that the ASAP mandate promotes changes toward high standards and constructivist education. Furthermore, we knew from the findings of the policy study (Noble, 1994) that the state had made no provision for capacity building in support of the mandate. Nor had the state attended to issues such as delivery standards or opportunity to learn. Thus, this mandate is unfunded and the provision of professional development provided by the state in support of change was meager or nonexistent. The only state mechanisms to instigate the change included the power of the ADE to persuade through rhetoric (e.g., repeated reminders to district officials and teachers of the importance of the Arizona Essential Skills and of teaching "the way kids learn"), the threat of disapproval of the District Assessment Plans, and the performance test itself (which was initially perceived to be low-stakes), plus the preliminary Forms A, B, and C and workshops to train teachers how to administer and score the assessment. Therefore, we recognized that both the resources for changing toward the promoted goals and the authority and power to change had to be understood at the local rather than state level. Based on these assumptions and understandings, we chose four categories

	VALOR	FRANKLIN	PINES	HILLDALE
RESOURCES FOR CHANGE	MATERIAL: Impoverished. KNOWLEDGE: No gatekeepers. Incompatible curriculum. No professional development.	MATERIAL: Adequate, but dissonant with ASAP. KNOWLEDGE: Scattered, marginalized. No gatekeepers. Incompatible curriculum and professional development.	MATERIAL: Rich, but dissonant with ASAP. KNOWLEDGE: No gatekeepers. Incompatible curriculum. Incompatible professional development.	MATERIAL: Rich, consonant with ASAP goals. KNOWLEDGE: Several knowledgeable gatekeepers. Professional development in sync with ASAP.
POWER TO CHANGE	Laissez-faire organization. School and teacher discretion.	Top-down power, centralized district, no teacher power.	Policy-driven, top-down, centralized, standardized. No teacher or principal power.	Site-based power, principals lead and persuade teachers. District encourages, supports. Teachers have power.
ASSUMPTIVE WORLDS	DISTRICT: Concrete-sequential. TEACHERS: Vary. Dominant image of pupil as disadvantaged/ deficit.	DISTRICT: Concrete-sequential. TEACHERS: Concrete-sequential or marginalized. Dominant image of pupil as disadvantaged/ deficit.	DISTRICT: Concrete-sequential image drives everything. TEACHERS: Concrete-sequential or marginalized.	DISTRICT: Constructivist. TEACHERS: Constructivist or marginalized.
ROLE OF TESTING	Low stakes, Low expectations, Low burden. ASAP is just add-on accountability requirement.	High stakes on district CRT. Low expectations. Rejection of ASAP as inappropriate for disadvantaged. ASAP adds to high test burden.	Low expectations. Very high stakes on district CRT. Very high test burden. ASAP adds to test burden.	High expectations, but moderate stakes from district. Low test burden. ASAP fits instruction, but still fills accountability function.
STATUS AT END OF YEAR ONE: REACTION	Teachers accommodate by "disintegrating" integrated tests and lessons. No capacity-building. Resistance: constructivist teacher resigns.	Passive acquiescence. Disintegrating. Added test burden creates frustration. Constructivist teachers isolated. No capacity-building.	Minimal, symbolic accommodation. Contrary philosophy still dominant. No capacity-building. Constructivist teacher resigns. Test-wiseness training.	Alignment of district scope and sequence to ASAP. Teachers vindicated by perceived match between their philosophy and ASAP. Test-wiseness training. Constructivism NOT CAUSED BY ASAP. Capacity-building.

Figure 1. Analytic matrix.

that seemed to account for the status of the site at the end of the first year. For example, the curriculum and pedagogy at Valor were virtually unchanged after one year of the program. No resources were available to direct toward ASAP-consonant activities, and thus no capacity was developed. School personnel acquiesced to the ASAP requirements, and ASAP merely added to the accountability load. Some resistance was evident in the departure of one of the constructivist teachers who experienced this burden. The status of change can be attributed in part to resources issues, knowledge, assumptive worlds, and organizational culture there.

The categories in the Analytic Matrix are listed and defined as follows.

- **Resources for Change:** *Material Resources* refers to the district's financial capacity to purchase or develop curriculum and to offer teachers professional development activities consistent with ASAP goals. Where financial resources are available, we ask whether they are directed at activities consonant or dissonant with ASAP aims. *Knowledge Resources* refers to the presence in the district and school of officials and teachers with knowledge and commitment to constructivist education and performance testing. Each site was characterized according to whether there was some gatekeeper, such as principal, coordinator, consultant, or other person who could interpret ASAP procedures and help teachers make changes consistent with ASAP aims. In some sites, a coordinator had been named by the district, but the person lacked knowledge, was unavailable to teachers, or soon left the district, and thus failed to help teachers make consonant changes.

- **Power to Change:** We characterized each site according to its organizational culture and where the power exists to make changes at the classroom level. For example, a centralized and hierarchical district vests control over change at the district level, leaving teachers and principals with little discretion to change in contrary directions. Local options remaining include acquiescence, accommodations (e.g., dis-integrating integrated curriculum or inappropriate test preparation), resistance, and marginalization.

- **Assumptive Worlds:** In this category we condensed the images of the pupil, learning, teacher, and curriculum that seemed to characterize both the district and the teachers at each site and the extent to which the dominant philosophy was either consonant (i.e., constructivist) or dissonant (behaviorist or concrete-sequential) with ASAP aims (assuming that ASAP is in fact constructivist). The constructivist assumptive world views the pupil as an active meaning-maker, the teacher as a coach or partner in meaning-making, and the curriculum as thematic, integrated, and negotiated, consistent with

pupil interests and prior knowledge. The concrete-sequential assumptive world views the pupil as an empty receptacle, teacher as conduit of curriculum and imparter of skills, and the curriculum as a hierarchical set of standard skills for the pupils to master.

- **Role of Testing:** This category reflects our characterization of the test demand or degree of testing stakes imposed on classrooms at each site. We distinguish the perceived function of tests as accountability devices (performed for external audiences) rather than as integral parts of instruction and whether there is a strong demand for high scores or measured change at the site. We also note the degree of test burden (proportion of time consumed by various testing functions), the expectations at the site for high or low scores based on past history, and where ASAP fits into the testing scheme.

- **Year-end Status:** This row in the matrix reflects our perspective of where each site stands with respect to reactions to the ASAP mandate.

Conclusions

What does the multiple case study tell us about the effects of the changing test mandate? Variations of local response are both substantial and significant. After a year of implementation, ASAP has largely been absorbed and subsumed under local, but apparently more salient, concerns. The goal of enhancing constructivist, integrated, "thinking" curriculum and pedagogy has been addressed most directly only in Hilldale Elementary, a school characterized as not only economically advantaged but also already well on its way toward ASAP goals before (or independent of) the state mandate. Many Hilldale teachers were already trained in and committed to holistic pedagogy, having, for example, a literature-based rather than a basal reading program and integrated, thematic curriculum. Its principal shares constructivist assumptions and acts as an agent of change at both the school and the district. The district administrators view performance assessment as "the wave of the future." They accept the district responsibility (given that the state had made no provision for it) for the professional development of teachers to gain the expertise that holistic, integrated, "thinking" education requires. The district had financial resources and aimed them toward acquisition of compatible materials and professional development. Teachers take courses, seek consultants' advice, and participate in staff development workshops, all consonant with ASAP aims. There seems to be a culture of teacher professionalism (time and a certain degree of autonomy—yet teachers who

hold on to behaviorist images are in the minority and marginalized) to advance constructivist curriculum and pedagogy. Several key teachers have made it their responsibility to serve on committees and take workshops related to ASAP and scoring rubrics. Although there is an intense interest in high scores among district officials and parents, the history of high test scores and awards at Hilldale provides Hilldale's principal with a degree of autonomy probably not experienced by every school in the district. Even so, the organizational culture in the district is decentralized, with power to make changes diffused among the schools. District administrators provide impetus to change through the power of persuasion and capacity building. The criterion-referenced testing program previously used in the district has been abandoned in favor of portfolios and ASAP Forms A, B, and C. Thus, a further barrier against change toward constructivist education has been removed. Nevertheless, even Hilldale teachers recognize that ASAP serves an accountability function, and they direct attention to the test as a test and to what aspects of the test will pay off in high scores.

Contrast these characteristics with those of Pines. Pines' economic resources are equal to those of Hilldale. Yet few resources were aimed at acquiring materials or training its teachers in support of ASAP aims. Teachers who themselves are supportive of those goals are torn between pursuing those goals and satisfying district requirements. The culture of the district is centralized and policy-driven, specifying almost every curricular decision. It is backed by a prescriptive, district criterion-referenced testing program and strong demand for high scores on those tests. Once district curricular and testing requirements are met, there is very little time and energy left for teachers to pursue alternative instruction. Teachers acquiesce to district images of curriculum, instruction, pupil characteristics, and assessment, or else resist by leaving the environment. Because Pines is a relatively low-scoring school in a district that scores high and whose administrators and parents insist on high scores, the accountability pressures are extreme. Teachers' evaluations and principal's positions are perceived to be on the line. In this set of circumstances, ASAP-related goals seem relatively remote and irrelevant to teachers' concerns. ASAP adds to the accountability burden. Teachers accommodate by dis-integrating and focusing attention on what will be scored, but make few changes toward constructivist

education. At the end of one year, little capacity has been created that could logically lead to authentic changes.

At Franklin, the degree of disadvantage of the pupil population is central to teachers' and administrators' rejection of ASAP. The official view of pupils is that they come to school as empty vessels that must be filled, a drop at a time, with skills. These skills are considered to be hierarchically arranged so that higher order thinking or problem solving can only be pursued once basic skills are mastered. Thus, ASAP, which requires integration of, for example, reading and writing with mathematical calculation, is viewed as beyond the reach of Franklin's pupils, who are almost exclusively poor, minority, and limited English-speaking. As at Pines the majority and official view silences the few teachers who might think differently. Materials and district tests that emphasize the concrete-sequential curriculum and behaviorist pedagogy sustain the dominant images. There is a high demand for demonstrated growth and high scores on the district criterion-referenced tests. The tests *are* the curriculum, in fact, in that there are very few instructional transactions outside the scope of the tests. District administrators define "master teachers" as those whose students get high scores. In a classic recreation of Taylorism, the principal designates master teachers to design instructional packets they have found to be successful for attaining high scores and dispenses those to the other teachers. District CRTs are constructed by teachers and considered to be the only measurement that suits this population—not ITBS and certainly not ASAP. Though Franklin has sufficient economic resources to modify its instruction and train its teachers toward ASAP aims, there is little chance that it will do so, so powerful is its culture to the contrary. By the end of the first year, the only changes evident are a passive acquiescence to the added accountability burden of ASAP and accommodation by dis-integrating and focusing on scores.

Valor matches Franklin in the degree of disadvantage of its pupils, yet its rural, agricultural economy impoverishes district resources, making modification of local curriculum, instruction, and teacher training problematic. Also like Franklin, its pupils score low on standardized tests, but in contrast, there is relatively little pressure on teachers to raise those scores, either on the state-mandated ITBS or on local criterion-referenced tests. Valor's organizational climate is laissez-faire. Curriculum decisions are not

driven centrally. Teachers have a degree of autonomy greater than the other three sites. The focal teachers observed in this site, therefore, varied among themselves in their images of pupil, teacher, curriculum, and assessment, some consistent and others inconsistent with ASAP-related aims. What overwhelmed culture and image at Valor, however, was the limitation in resources. Textbooks were twenty years old and incompatible with ASAP. The district's purchase in the 1980s of an off-the-shelf criterion-referenced testing program (in format similar to the ITBS) represented such a substantial investment that it is unlikely to afford a new one, more fitted with process and integrated performance assessment. There was no individual in the school or district who could interpret state images or inform teachers about what needed to be done to adapt to ASAP. When the time came to administer ASAP, teachers struggled with its complexities, showing most clearly how teachers' prior knowledge of an instructional activity must be taken into account if a mandating agency expects that activity to succeed. Valor teachers, though competent to teach what was familiar to them, had never experienced the use of writing to teach reading, for example, or how to teach estimation in mathematics by referring to familiar objects. The aim of "thinking education" is pupil understanding and integration of new knowledge by referring to prior knowledge. The aim of concrete-sequential education, is to repeat an activity until the pupils "get it right," as opposed to "getting it." But the Valor teachers' own prior knowledge was an inadequate scaffold to hold the holistic, integrated teaching, learning, and assessment model promoted by ASAP. When presented with an integrated unit in Form A or in the new social studies text the district adopted, the teachers actually "dis-integrated" it. That is, they decomposed the lesson into bits that they thought could be taught in such a way that all the pupils could perform correctly and get the right answer. There was no money for in-service training that might have helped the teachers make the change.

What was common among the sites was the belief that testing that comes from an outside agency is still testing, with its attendant considerations that testing must be individualistic, competitive, silent, and objectively scored. Testing that is done for outside agencies is still separate from instruction and added on to normal school activities. Assessment from the teachers' point of view is what advances instruction day-to-day, for which they have multiple

indicators besides test results. This view of ASAP as an add-on, done to satisfy an external audience, contradicts the state policy image that ASAP testing should be integrated with instruction rather than be a supplement. The state image implies that local curriculum not consistent with the Arizona Essential Skills and ASAP assessment should wither away. But teachers and principals orient themselves more to local demands and see state requirements as an unwarranted intrusion or unlikely to persist.

Across the sites, teachers viewed ASAP as low-stakes and aimed more to change instruction than to evaluate the efficacy of schools. Some even regarded ASAP as a pilot or experiment. The exception was at Hilldale, where teachers were more knowledgeable about how the ADE intended to conduct the on-demand Form D assessment and about the scoring rubrics that would be applied to the performance test results. The focal teacher, Terri, who was the ASAP liaison for Hilldale, directed the attention of her pupils to those features of the performance assessment that would, in fact, be scored and the attention of her colleagues to the accountability function that ASAP was likely to serve.

The view of ASAP as a low-stakes test designed to nudge districts toward a different form of pedagogy was overturned when, in the spring of 1993, the ADE reported ASAP scores by school and grade level, in the same manner that it usually reports ITBS results. At that point, more practitioners and administrators viewed ASAP as part of the state's accountability package. By that point as well, districts began struggling with the notion of setting a cut-off score on the performance test that would demonstrate to the ADE the districts' mastery of the Essential Skills measured by each assessment.

This study has shown how the actions of practitioners are far from uniform in response to a policy mandate. Local interpretations and organizational norms intervened to color, distort, delay, enhance, or thwart the intentions of the policy and the policy-shaping community.

It is, however, only the story of the first year of implementation of a measurement-driven reform, under perceived low-stakes conditions. The proponents of such reform might be heartened by the prospects of change under conditions of increased stakes, brought along by the ADE's publicizing school scores, attaching mastery levels to the performance test, and attempting to make grade promotion and high school graduation related to

performance on ASAP. Such ratcheting of stakes may increase educators' attention to changing instruction in the desired direction. Or, the reaction may be to do what is necessary to increase the scores themselves, as the literature on dysfunctional side effects of accountability suggests (Campbell, 1979).

In either case, the prospects for reform toward the aims of the mandate must be judged in light of one notable barrier, the variable status of teachers' expertise or prior knowledge of holistic, integrated, thinking curriculum and pedagogy. Hilldale teachers have reported that it took years of expert guidance, and time to experiment, reflect, and collaborate, once they personally made the commitment to change in this direction. No institutional obstacle was placed in their path. The distance on this dimension between Hilldale and the other schools we studied is vast. The means for schools to traverse this distance have been ignored in policy formation and administration, or left to the vagaries of district and school practice.

References

Blumer, H. (1986). *Symbolic interactionism: Perspective and method.* Berkeley: University of California Press.

Campbell, D.T. (1979). Assessing the impact of planned change. *Evaluation and Program Planning, 2,* 67-90.

Erickson, F.E. (1986). Qualitative methods in research on teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119-161). New York: Macmillan.

Gold, R.L. (1958). Roles in sociological field observations. *Social Forces, 36,* 217-223.

McDonnell, L.M., & Elmore, R.F. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis, 9*(2), 133-152.

Miles, M.B., & Huberman, A.M. (1984). *Qualitative data analysis: A sourcebook of new methods.* Beverly Hills, CA: Sage Publications.

Noble, A.J. (April, 1994). *Measurement-driven reform: Research on policy, practice, repercussion.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Noble, A.J. & Smith, M.L. (in press). Old and new beliefs about measurement-driven reform. *Educational Policy.*

Rein, M. (1983). *From policy to practice.* Armonk, NY: M.E. Sharpe, Inc.

Resnick, L.B. (1989). Toward the thinking curriculum: An overview. In L.B. Resnick & L.E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research 1989 ASCD Yearbook* (pp. 1-18). Reston, VA: Association for Supervision and Curriculum Development.

Smith, M.L. (1993). *The role of high-stakes testing in school reform.* Washington, DC: National Education Association.

Smith, M.L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1990). *The role of testing in elementary schools* (CSE Tech. Rep. No. 321). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.

Smith, M.L., Noble, A.J., Cabay, M., Heinecke, W., Junker, M.S., & Saffron, Y. (1994). *Case studies in measurement-driven reform* (Report to CRESST, OERI Grant No. R117G10027, CFDA Cat. No. 84.117G). Los Angeles: University of California, Center for Research on Evaluation Standards, and Student Testing.

Weatherly, R. & Lipsky, M. (1978). Street-level bureaucrats and institutional innovation: Implementing special education reform. *Harvard Educational Review, 47*(2), 171-197.

APPENDIX

CROSS-SITE DATA MATRIX

CROSS-SITE DATA MATRIX

	VALOR	FRANKLIN	PINES	HILLDALE
DISTRICT CHARAC-TERISTICS	Unified K-12 [K-5, 6-8,9-12]. 1000 pupils. Dropout rate 12%. Transience rate*, fluctuates according to harvest seasons. Rural site, migrant influence. District and school are ethnically diverse.	Elementary, K-8 [K-2, 3-8]. 800-900 pupils. Dropout rate from high school district is 50%. Transience rate 90% Urban core. District and school are ethnically diverse.	Elementary, K-8. 10K pupils in district. 14 Schools are K-5 and 4 are 6-8. Dropout rate from high school district is 16%. Transience rate*. Suburban. District is mostly middle class and Caucasian (17% ethnic minority).	Unified, K-12. Dropout rate is 10%. 24,000 pupils in district. Less than 10% minority pupils. Suburban.
DISTRICT RESOURCE BASE	One of the most resource-poor districts in AZ agricultural, commercial economy. Almost no discretionary funds are available for curriculum or professional development.	Nearby factories provide tax-base adequate for financing school operations. Moderate resource base.	Residential, commercial, light industry provide adequate tax base for financing school operations. Moderate to good resource base. District invests substantial amounts in professional development.	Residential, commercial. Adequate tax base for financing school construction and programs.

	VALOR	FRANKLIN	PINES	HILLDALE
DISTRICT TESTING MODEL	ITBS in Grades 1-8, due to Chapter 1 requirements. Purchased off-the-shelf criterion-referenced test package, similar to ITBS format, given twice/year. CRTs are objectives-referenced with 80% mastery level set. Administrator: "85% of students' composite scores (ITBS) fall below the 50th percentile." District pays for copying costs and costs of hiring teachers to score the ASAP Form A for Grades 9-11 to conform to state mandates for district testing. This is considered a significant expense. District feels state sends conflicting messages: that district tests too much AND that they should expand performance testing. "But, hell, give us a break."	ITBS in Grades 2-8, due to Chapter 1 requirements. District CRTs written by teachers, given 3/year with 70% passing rate specified. Teachers write CRT items, which resemble ITBS format plus writing test. District test results used for teacher evaluation and curriculum revision. Always one of the lowest scoring districts on ITBS. Administrator: "2 or 3 standard deviations from the state norm." Although district claims to "totally deemphasize ITBS," it publishes scores in newsletter to parents.	ITBS in mandated grades plus Grades 5 & 6. District CRTs developed by district testing bureau, administered 2/year in math and communication arts. Format is standardized, objectives-referenced, multiple-choice with mastery levels specified; timed math tests. Frequent formative tests are expected, with results recorded by teachers and monitored by district. District test results used for teacher and principal evaluation and school comparisons. District claims disinterest in ITBS results, but publishes them in information packet, showing district is at or well above national norms in all tested areas. CRT results used in promotion/retention decisions.	ITBS in mandated grades. District testing went through revision post-ASAP from CRT to ASAP Preliminary Forms A, B, and C (performance assessment). No testing bureaucracy exists at district. Little commitment to a standardized testing model.

	VALOR	FRANKLIN	PINES	HILLDALE
TEST DEMAND	Although "there's a big emphasis on test scores," district does not expect high scores, based on its history and pupil composition. Although gain scores are perceived to be accurate indicators of how teachers are doing, administrators try not to focus teachers' attention on low scores or pressure them to increase scores. Teachers don't feel too much pressure, regardless of which test is used.	District' mission statement's first goal is to raise pupil achievement as measured by CRT. Teachers are compared on CRT scores, are told to "fix" whatever skills their students are low on. Teacher: "Administrators will do anything to make those scores look better," even changing items in midyear. Total de-emphasis on ITBS score. High stakes on CRT only.	The high-stakes testing environment is an extension of its philosophy of learning, which emphasizes measured mastery of skills. District statement of philosophy: "Schools can maximize the learning conditions for all students through clearly stated outcomes, high expectations for all students, and continuous assessment of student learning." There is a strong perception that the schools exist on a distribution of their test scores. Principals of low-scoring schools are publicly admonished and subject to transfer or removal. District test coordinator to principal: "Pete, is there any life at Pines?" Teachers, in turn, feel the pressure exerted by principals. Anxiety is translated into rigid compliance to district testing policy. CRT results are used in part in the district career ladder program. Use of CRT results for promotion decisions indicates high stakes for pupils. Media attention to test scores is strong.	District administrators expect scores to be high because of socio-economic status of residents and because standardized test scores have been historically high. District projects a positive image in part because of its high-ranking scores. Principals of low-scoring schools recognize this competitiveness, yet pressures are not particularly high to improve scores at any cost. District administrators are more interested in authentic improvement rather than in high scores for their own sake. Reaction to ASAP scores that were not as relatively high as ITBS/TAP scores, however, prompted district official to say to principals: "We should not be below the state or county average. I intend to be involved with you so the 1994 results show a different response." Principal in the school with highest ranking ITBS scores was concerned that ASAP school scores were not at the top, but did not encounter pressure from district officials to raise them. Strong media attention to scores of all kinds.

BEST COPY AVAILABLE

	VALOR	FRANKLIN	PINES	HILLDALE
DISTRICT ORGANI-ZATIONAL CULTURE	District is very loosely organized by tradition and rural character. Principal has relatively wide range of autonomy from district administrators, and in turn imposes few mechanisms that would withhold authority from teachers. No bureaucracy exists for either curriculum or assessment.	Mission statement: "Increased emphasis will be placed on instructional and management strategies which enhance improved academic performance." Top-down authority structure, with CRTs used as standardizing, centralizing mechanism.	District document: District is actively engaged in efforts to improve teaching, learning, and decision making through systematic efforts (Total Quality Management). Administrator: "This is a policy-driven district." While espousing site-based decision making, has developed elaborate procedures for monitoring the performance of schools, principals, and teachers and pushing a common, centrally-controlled set of goals and operations. District continues to experience rapid growth and frequent changes in administrative personnel. District testing plays a dominant role in the top-down nature of the district, with principals defined as responsible for producing good test scores and schools judged as good or bad based on scores. Administrator: District is driven by pattern of promoting and adopting whatever is perceived to be cutting-edge educational movement, which keeps teachers always in the role of novice, without time or resources to become experts.	There is a traditional value placed on school authority balancing central district authority. District is experimenting with site-based management. Schools differ from each other on the reading programs they offer, their graded organizational structure, etc. District officials lead by encouragement and capacity building, rather than by strict policy mandates or intimidation. Official: "It's up to the district to provide for staff development so that teachers can respond to innovations such as ASAP. We have the resources and intend to make them available to teachers."

	VALOR	FRANKLIN	PINES	HILLDALE
ASAP EXPERTISE/ COMMIT-MENT OF DISTRICT	District officials accept performance assessment and holistic instruction as an ideal, but lack in-depth understanding and expertise to pass on to teachers. Misunderstandings are evidenced, for example, by claiming that whole language is an emphasis in the district, but the adopted reading program follows a direct instruction model. Officials fail to recognize contradictions such as these.	Commitment of administrators to basic skills mastery precludes commitment to ASAP-like teaching. Administrator: "Local determination of ... what needs to be fixed with a kid academically always supersedes what the legislature says." No ASAP gatekeeper exists. Prior failure of holistic instruction experiment is additional barrier.	A significant and vocal minority of teachers profess commitment to ASAP principles of teaching and assessment. Most district administrators, however, are more interested in contrary principles, such as those embodied in TQM. Key administrator thinks that performance assessment is "the performance of objectives." No district official was expert in ASAP or advocated for it.	Expertise in ASAP and ASAP-consistent education was scattered across district, not centrally situated. ASAP coordinator was part-time only and was replaced after the year. Teacher (speaking about coordinator): "She's just a gifted ed teacher, so what good is she?"
BELIEF OF OFFICIALS IN ASAP PERMA-NENCE	Officials believe that ASAP has merit as a form of assessment, but are concerned about its costs, feasibility, and its possible use by the state as an accountability mechanism. They accept it as a permanent part of the state's accountability package and generally accommodate.	Administrators perceive ASAP to be temporary and of little value in shaping education for "these pupils." "It's going to collapse logistically after 2-3 years." Thought by some to be a "pilot."	Administrators believe that ASAP has problems but will probably be a permanent part of state test mandate, and will work out the kinks. But they see ASAP as much more a mastery model than a constructivist model.	District official, referring to performance assessment: "This is the wave of the future. This is reform." Official speaking to principals: "For those who believe this type of assessment is going away—it ain't going to happen." Teachers fear that ASAP will be like many movements in AZ education, here and gone, in relation to the career interests of state officials.

	VALOR	FRANKLIN	PINES	HILLDALE
REACTION TO ASAP RESULTS	District was above average on ASAP results (attributed by district officials to a fit between "what we teach and AZ Essential Skills"). This contrasts with their much lower average performance on ITBS. District test coordinator presented results at public school board hearing, expressing pride that Valor was second highest in the county. Results were not translated into any action, however.	District ASAP scores were only 1 standard deviation below state average (not as relatively low as ITBS scores). Little notice was taken. Some acknowledgment that state emphasis on ASAP might require some district adjustments. Third-grade teachers' rewrite of test items to reflect ASAP principles was reversed in favor of more basic skills.	ASAP results were not given to schools in any systematic or purposeful way. Most teachers did not even know about them. Because the school scored higher in the district ranking of ASAP than it typically scores on ITBS or CRT, the principal felt some pride and received a few congratulations. However, because ASAP was given no importance by the district, this accomplishment was greatly overshadowed by the school's performance on the more valued district tests.	Official: "We should not be below the state or county average. I intend to be involved with you so the 1994 results show a different response. It is time to get on board. Make sure those in your building understand that alternative assessment is an important tool for teaching and assessment. If your students have not been exposed to higher order thinking then they will not be prepared to do these things." Given the small variance of ASAP scores across schools, this statement was an over-interpretation and undue cause for concern. Principal was perplexed at Hilldale not being the highest ranked school, since ASAP was closest to "what we teach."

	VALOR	FRANKLIN	PINES	HILLDALE
PROSPECTS FOR SECOND YEAR CHANGE	Given the laissez-faire district culture, some incremental changes are possible among teachers who are already predisposed to holistic instruction and performance assessment, but little change should be expected among teachers with other predispositions. No strong organizational values act as barriers to change. Neither are there supportive values or mechanisms. The lack of resources for curriculum and professional development also mitigate against change in the immediate future.	The inherent conflicts between ASAP principles and local beliefs about the nature of pupils, etc. makes change in year 2 unlikely. Some accommodation may occur with great cognitive dissonance. After ASAP scores came out, administrators conceded that they might have to pay more attention to ASAP scores.	Support for ASAP-consistent instruction and assessment are overwhelmed by the top-down organizational culture, tightly structured connection between curriculum, testing, and management philosophy, strong testing bureaucracy, absence of time and support for alternatives. Incremental changes may occur, such as incorporating Forms A-C in district assessment of writing, yet wholesale paradigm shift will have to occur to make significant changes.	District goals, organizational culture, testing program are consistent with the ASAP reform, though not as a consequence of that reform. Resources are available to continue to develop teachers and curriculum. Scope and sequence and text book adoption have been aligned with ASAP. A critical mass of teachers and administrators exists that have training in whole language and conceptual math. These factors suggest that prospects are positive for second year movements toward ASAP goals. Negative foreshadowing concerning the concentration by teachers on those aspects of ASAP that are scored, and possible narrowing of focus toward them.

	VALOR	FRANKLIN	PINES	HILLDALE
DISTRICT IMAGE OF PUPIL	The most salient image is that pupils with economic disadvantages have DEFICITS that schools must address and that hold back progress toward higher order, thinking instruction. Students are perceived as having levels of ability and motivation that exist independent of teaching efforts. "I don't understand the standard for higher order thinking skill. The student down here is thinking at the highest level he can possibly think."	Pupils who are highly disadvantaged are unlikely to benefit from reform efforts. Disadvantage is the driving force that shapes goals and activities, including the emphasis on test scores on basic skills. Low scores are attributed to poverty, an enduring quality irrespective of teacher efforts. Pupils are empty receptacles when they come to school. Pupils here can't transfer or generalize.	Pupils have particular levels of learning needs that must be diagnosed, then must be given sufficient opportunities to practice the skill until it is mastered. Focus is accurate mastery rather than understanding. Learning is extrinsically motivated. Convergent learning is more important than divergent. If skills are not repeated, they will be forgotten.	Pupils bring with them a variety of knowledge, interests, and capabilities that must be considered in designing curriculum and school organization, by those closest to them. Pupils are active meaning-makers. Pupils need thinking and problem-solving skills to help them succeed in the world. Pupils' social and psychological needs are as important as their academic needs.
DISTRICT IMAGE OF TEACHER	Officials believe that teachers should all be following common objectives and materials but impose few mechanisms to insure that. They believe that low test scores are sometimes good for teachers in that they provide chances for self-evaluation and revision. Teachers are generally on their own, for better or for worse. As a result, teachers vary, and their own beliefs and images about pupils, teaching, and curriculum are more salient than district images.	Each teacher's scores are public and teachers with low scores are told to just "fix it." Teachers are evaluated by scores. Teachers with highest CRT scores are designated as "master teachers." They are given release time to develop skill-building activities for other teachers to use.	Teachers are the transmitters of curriculum. Their role is to diagnose "the correct level of difficulty" each pupil needs in relation to a required curriculum component and devise strategies to make pupils successful at that level. Teachers provide subskills practice in anticipation of district testing. Teachers regarded as emotionally overreacting to test emphasis in district; may not have enough competence to teach higher order thinking curriculum.	Teachers ultimately govern what goes on in classrooms, and teachers will change toward ASAP style instruction only if their philosophies are consistent with it. Change comes through building teachers' knowledge and demonstrating what can work. Teachers can be leaders in supporting the changes that other teachers should make. ASAP mandate can serve as a stimulus or guide to provoke teachers' experimentation and reflection. Ideally, teachers provide instructional opportunities and materials for pupils to use in constructing meaning.

	VALOR	FRANKLIN	PINES	HILLDALE
DISTRICT IMAGE OF CURRICULUM	District officials would prefer a curriculum that is standardized and one that results in an accumulation of specific skills. "Without a district curriculum, teachers get married to a text" and don't think beyond it to a coherent curriculum. Curriculum should be aligned to common state and district objectives and tests. Five years earlier, preferences for whole language teaching were explored through a series of staff development activities. But momentum was lost and no funds were available to change texts and district tests. Most teachers ignore or are unaware of district curriculum. No textbook adoption procedures. Principal feels curriculum is district's weakest area.	There is no curriculum other than what is on the CRTs. Principal: "If you're teaching something that isn't there, either quit teaching it, or else" write an item. Basic skills must be mastered before thinking skills. Constant repetition and drill are considered essential. Skills are taught separately, not integrated. Reading instruction is highly tracked. Progress through grades influenced by scores.	Curriculum guide is centerpiece for controlling what goes on in schools. A pie chart specifies how much instructional time should be spent in each of 12 curriculum components. 59% is to be spent in reading and communication arts (the components covered by CRTs) at the third grade. The curriculum is packed, leaving little room for divergence. Skills are hierarchically arranged. Even a reading series with "Integrated" in its title is used by fragmenting literacy into separate subskills and practicing those to mastery levels.	AZ Essential Skills plus district objectives provide a framework within which teachers select appropriate materials and instructional opportunities. Writing process is pervasive across curricular areas. Schools choose texts and curriculum packages. Curriculum should be integrated because that is what is meaningful to students (although secondary teachers resist the idea). Some key district officials subscribe to constructivist ideals. The message to principals that, although district has adopted Open Court (direct instruction), schools should do whatever makes sense and succeeds for their particular population.

	VALOR	FRANKLIN	PINES	HILLDALE
DISTRICT IMAGE OF ASSESS-MENT	Officials believe that tests should be aligned with local curriculum to be valid. ASAP fits this criterion, but is problematic because of its costs and feasibility and because the state does not mandate (or pay for) administration of ASAP at each grade level. "The concept is great but the feasibility is humongous." They believe that assessment should provide comparative data, but also should reflect "what we teach" at every grade level; should also show mastery of basic skills but also diagnose pupil needs for teachers' use; yet should be inexpensive and feasible.	Externally mandated tests are culturally and linguistically biased. Only locally constructed tests are valid for "our population." Testing and curriculum must be matched because pupils can't transfer or generalize. Teachers are told to do whatever it takes to raise CRT scores.	Assessment is for continuous monitoring of progress and is inextricably linked to curriculum. Document: "The foremost purpose of district-wide testing of ... objectives is to evaluate and monitor student growth, document that growth ... help the District evaluate instructional program effectiveness, and ... partially evaluate teacher effectiveness." A district administrator has sole responsibility for the testing program and actively and effectively promotes that agenda. Administrator: "Assessment must be tied to objectives, systematic and objective." Validity is content validity and demonstrated pre-post gain (sensitivity to instruction). Teacher accounts aren't credible for making system-wide conclusions and decisions. ASAP also not likely to meet that standard.	Officials believe that instruction and assessment should be integrated and ASAP supports this. Test results reflect in part the socio-economic composition of the community. Still, test results should be high to show the district or school is doing well. Test scores play only a small part in teacher or principal evaluation, but a pattern of low scores might indicate need to improve instruction. High ASAP scores would show that districts have changed in the right direction; i.e., toward holistic, real-world, problem-solving education (consequential validity). But, officials reserve judgment on ASAP, waiting to see how the state will report or use scores. Ranking schools by ASAP scores or using scores to indict teachers will alter reactions and encourage teaching toward the test in possibly inappropriate ways.
COMMUNITY CHARAC-TERISTICS (SIZE AND TYPE)	School is located in an unincorporated town of 600 residents, within 100 miles of Phoenix, near an Indian reservation. Economy is predominantly agricultural.	School is located in urban core. A few small, well-kept homes are near government projects, shacks, and homeless shelters.	School is located in a middle class suburb, amidst well-kept tract homes. The school draws from an attendance area that is much more economically and ethnically diverse than the district average.	School is located in a prosperous, middle- to upper-middle-class urban/suburban community.

	VALOR	FRANKLIN	PINES	HILLDALE
COMMUNITY SES	Community is predominantly lower-middle to lower class. Unemployment rates are high due to fluctuation in agricultural industry.	One of the most impoverished communities in AZ.	School serves a mixture of middle-class and upper-middle (professional and management), as well as a poor enclave of poorer, minority "apartment" dwellers.	School serves middle -to upper-middle-class families, most of whom believe that "their children are all gifted."
PARENT INTEREST AND PARTICIPATION	Parents want to know how students compare with students in the rest of state and nation. "There's a big interest in test scores." Parents rarely question school's authority. Many are undereducated and unfamiliar with the workings of school. Parents want to know how students perform because they accept the authority of the test scores.	About 25% of district families take an active interest in the school and participate in school activities. Majority are only minimally involved. Teachers: "Parents have little interest in test scores here."	The upper-middle-class parents are actively involved in school life and intensely interested in test scores. Principal and experimental reading program designed to raise low scores were the focal point of conflict during year of study. Principal: "Parents came breaking down doors, with newspaper articles in their hands, saying what's the deal? Why are we always lowest? And we had to respond to that and make efforts to make the scores look better."	Parents participate in school activities at high rates and are extremely interested in test scores. Principal, however, plays down standardized test scores and actively tries to educate parents on alternative ways of indicating achievement.
PUPIL SES	75% of pupils at school qualify for free/reduced lunch. Many migrant families from Mexico send children to school here during harvest.	97% of pupils qualify for free/ reduced lunch Administrator: "The most disadvantaged, at-risk students in the state are at this school."	22% of pupils at school qualify for free/reduced lunch program.	School does not participate in reduced price lunch program.
PUPIL ETHNIC COMPOSITION	Anglo 41%, Hispanic 35%, Native-American 24%.	Anglo 8%, Hispanic 84%, Native American 2%, Black 6%.	Anglo 72%, Hispanic 20%, Native American 2%, Black 4%, (Unknown 2%).	Less than 5% non-white.

BEST COPY AVAILABLE

	VALOR	FRANKLIN	PINES	HILLDALE
PUPIL LANGUAGE DOMINANCE	Most Native American and Hispanic students come to school with English proficiency because of reservation preschool program or Valor preschool. About 35% of pupils are classified LEP.	Over two-thirds of pupils have first language other than English. One-half enter kindergarten with little or no English proficiency.	5% of school's pupils are limited English proficient.	Less than 5% have first languages other than English.
SCHOOL ORGANIZATION AND SIZE	K-5 plus federally-funded preschool. 500 students	Grades 3-8 plus Head Start, migrant education, transition English programs. 550 pupils.	Grades prekindergarten to 5. 550 students. Transience rate of 7%.	K-5. 900 pupils.
SCHOOL STRUCTURE	Traditional graded structure. Community distribution center provides food and clothing to community. Pull-out programs for Chapter 1 and computer literacy. Class sizes 25-32. Classes are heterogenously grouped. School policies discourage retention and few teachers practice it except in cases of extreme absenteeism.	Traditional graded structure. Class size averages 22-28. Chapter 1-eligible pupils pulled out for remediation. Retention based on CRT mastery (but attendance and teacher judgment enter in decision); little retention occurs. Instructional aides in all classes Transition classes and diagnostic/counseling services provided to transient students. Community Education classes provided in English literacy, computer ed, family math, and parenting.	Graded structure, but with several two-grade, combined classes. In each class, pupils are grouped for instruction and regrouped periodically based on CRT results. Promotion/retention decisions partially based on mastery test results. Chapter 1 school.	School experiments with various multiage and combined grade (3/4, 1-2) combinations. Enhanced fine arts, physical education and computer lab. Reading Recovery program available, but few are referred.

	VALOR	FRANKLIN	PINES	HILLDALE
GRADE-LEVEL CONFLICT	The disparities among classrooms in accommodation to ASAP are due to differences in teacher beliefs rather than grade-level differences. Third-grade teachers experience a greater burden of testing on instructional time than other grades. First- and second-grade teachers are interested in ASAP but express no sense of responsibility for it.	Third-grade teachers feel victimized by the need to satisfy conflicting demands and instructional principles (accountable for ASAP and CRT mastery). They feel that rest of school culture is discordant and nonsupportive. When third-grade teachers revised CRTs to be more compatible with ASAP, fourth-grade teachers reversed the changes, making them even more emphasizing basic skills.	Third-grade teachers tend to dismiss the importance of ASAP because it is not administered at every grade. Teacher: "We tried to get the second-grade teachers to try it, but they rejected it because it wasn't required." Third-grade teachers feel the test demands on them are doubled, because of ASAP mandate, or would be if district cared about its results.	Each grade level above K does some ASAP administration. Third-grade teacher: "It took a lot of pressure off us knowing that it wasn't just our responsibility."
PRINCIPAL ROLE	Principal plays a laissez-faire role with respect to ASAP, curriculum, etc. Sees her role as one of support and encouragement but, in rare instances, feels she has to coerce some teachers to perform. Believes in Hunter's Elements of Effective Instruction model.	Role of principal is to centralize authority over curriculum and teaching methods by evaluating teachers based on CRT scores, aligning instruction and assessment, and driving out untested content. A behaviorist with respect to both pupils and teachers—reinforces correct behaviors; exposes and shames incorrect behaviors; ignores process.	Principal is beleaguered because of the relatively low position of his school in the district distribution of scores. He serves as a conduit in a top-down hierarchy from district to teachers, conveying exactly the demands and expectations of the district. He has so little room to maneuver that he ignores ASAP altogether, along with any other goals and activities not specifically endorsed by district. He frequently reminds teachers of the importance of CRTs. Defines ASAP as alien to Pines teachers: "Teacher is teaching one method or style but testing method is another way."	Principal is a catalyst for change in directions ASAP seems to support. She can "play the testing game," but prefers to de-emphasize inflating scores in favor of authentically changing instruction and curriculum and developing teachers as professionals. Because she believes ASAP is valid as an indicator of constructivist teaching, she was bemused that Hilldale did not score at the top of the district. She is extremely knowledgeable about constructivist teaching and performance assessment and presses the teachers to get on board.

	VALOR	FRANKLIN	PINES	HILLDALE
ACCOMMO-DATION/ RESISTANCE OF PRINCIPAL	Principal accommodates to ASAP in the same way as to any other external mandate, with passivity.	Ignores and resists ASAP reform as irrelevant to local population and inconsistent with beliefs that basic skills must be mastered before thinking skills can be taught.	Principal recognizes that, in the current district culture, ignoring ASAP is safe.	Principal not only accommodates but acts as resource to the district in promoting activities consistent with ASAP.
TEST BURDEN	Chapter 1 requirements mean that ITBS is given to every pupil. CRTs given early and late in year. Sixteen days per year are spent in administering tests in third grade. Preparation time varies among teachers.	CRTs given 3 times/year with preparation time before hand. ASAP adds to third-grade burden. ITBS in all grades. Six weeks of instructional time taken up in testing in non-ASAP grades, eight weeks for Grade 3.	Teachers sometimes feel that all they are doing is testing, in one form or another.	ASAP is the district test at all levels but K, therefore test burden is light. ITBS only at mandated grades. Some teachers give CRTs at their discretion. District requires (and provides scoring for) an analogies test for determining academic ability.
TEST PREPARA-TION	Amount of preparation varies by teacher, based on perceptions of stakes. All but one teacher perceived ASAP as low-stakes and therefore didn't prepare pupils for it. One teacher had been burned by accountability function of tests in the past and therefore prepped for it.	Since curriculum and teaching are synonymous with CRTs, test preparation is constant. As testing schedule nears, teachers become more anxious and prepare more. Practice consists of timed tests and worksheets that use formats similar to CRT format. This is a daily event in some classes. Practice for ITBS is rare. Little practice ASAP. Form A stayed in box until week before D was to be given.	Since teaching and CRTs are part of the same package, test preparation for CRTs is on-going, becoming nearly frenzied as the spring post-testing schedule approached. 50% of instructional time is spent in test preparation. Preparation for ASAP was minimal. Some teachers didn't even "open the box" containing ASAP practice materials until the week before testing schedule. Focal teacher taught test-wiseness using Form A for 2 weeks prior to test schedule and administered ASAP-like exercises in form of homework. Emphasized ways of getting high rubric scores.	Principal has discouraged staff from using Scoring High. ASAP preparation consists of doing "activities" related to Forms A, B, and C.

	VALOR	FRANKLIN	PINES	HILLDALE
CURRICULUM/TEXTS	SRA text series (direct instruction model) is adopted reading text, although that is used only by some teachers. Those with constructivist tendencies embed more literature and de-emphasize the text. Most staff are unaware that a district curriculum even exists.	Curriculum IS the CRTs, said to also reflect AZ Essential Skills. Teachers follow Madeline Hunter's Critical Teaching Skills and have developed structured lesson plans for all skills. Reading text selected for match with CRTs, phonics-oriented basal series. Math text is available, but less important than CRT.	District specifies curriculum in 12 areas, separated, not integrated. Reading program, Success for All, has been substituted at Pines to raise low scores. Phonics-based, oriented to skills and some low-level comprehension. Math text is Addison Wesley (skills-oriented). Everyday Math is a less preferred alternative text that emphasizes problem-solving and writing about it. Text for social studies takes backseat to curriculum "bible." Science has taken a "backseat."	Reading program is literature-based. No basals are used, except by three teachers. Writing process unifies curricular areas. Math Their Way and Math a Way of Thinking. Although science and social studies texts are available, teachers tend to incorporate material in thematic units and projects. District provides materials for "hands-on" math and science activities.
TEACHER EXPERIENCE	Two focal teachers with 5 and 3 years of experience. Another had more than 20, with two master's degrees -and intentions to become administrator. District has tendency to hire new teachers who can't get into more desirable assignments. After a couple years experience, they leave. Teachers express pride in themselves as professionals and participants in the community.	Focal teachers had 16 and 30 years' experience. All teachers certified ESL as well as elementary.	Fourth-grade focal teacher has 15 years of experience and a penchant for technological innovations, particularly computers. Third-grade teacher has 6 years experience and interest in "total quality learning."	Focal teacher had 4 years' experience, all with same principal and constructivist training and continual professional development.

	VALOR	FRANKLIN	PINES	HILLDALE
TEACHER COMMIT-MENT TO HOLISTIC THINKING INSTRUC-TION	Because of decentralized structure, teachers varied on all teacher dimensions. Teachers fell roughly into two clusters: concrete-sequential (i.e., behaviorist, skills-oriented) and constructivist (e.g., whole language). Constructivist teachers were moderately committed, but even these have had limited exposure. Commitment a function of limited exposure and knowledge.	Those few who lean toward holistic instruction are viewed as outcasts. Emphasis by school on bit-by-bit instruction of basic skills and CRT content precludes integrated activities or thinking education.	Teachers see themselves as supporting the concept of holistic teaching, yet experience dissonance between the two paradigms. They redefine holistic instruction so as to make it compatible with a skills approach. Teachers engage in "dis-integrating." Teachers called ASAP a "round-about kind of learning." Teachers who favor constructivism rarely get the time or support to put these ideas in practice.	Teacher regards herself as a whole language teacher. Experts in whole language refer to the school as "in transition," i.e., having the rhetoric and moving toward thorough knowledge of holistic principles, but not yet deeply experienced or expert. Her reading program is literature-based and uses literature logs and literature studies. However, the interaction in groups is teacher-directed and focuses on convergent responses (e.g., "How does [author] use adjectives?" Has participated in several years' training in writing process and thematic teaching and is expert in these areas. Constructivist math expertise seems limited to thorough training and several years' experience in Math Their Way.

	VALOR	FRANKLIN	PINES	HILLDALE
PRIOR KNOWLEDGE OF ASAP-LIKE INSTRUCTION	The behaviorist teachers have none. Constructivist teachers had been practicing some form of holistic instruction in language and literature for about five years, since a series of workshops and training activities were held. But even constructivist teachers fell back to a basic skills approach to teaching math, because of inadequate scaffolding in the conceptual principles of mathematics.	Some Grades 1 and 2 teachers express some awareness, but only one is expert. "It took me ten years to [get to] the point of being comfortable with whole language, and I'm not as good as some." Prior knowledge is actually negative, as an earlier experiment with whole language instruction failed to raise test scores so was abandoned. Teacher: "We tried it and it didn't work here."	Teachers who claim knowledge actually misunderstand the principles of constructivist teaching; e.g., thinking that process-teaching and assessment cannot be grasped by low-ability pupils. Focal teacher shows she understands integrated, real-world teaching when she teaches a science lesson on owls, but doesn't translate that into teaching reading and writing , because latter is governed by CRTs.	Teachers have had workshops and college courses "for years" on writing process and literature-based reading. Some professional development has taken place in math (use of manipulatives to teach number concepts), but not nearly so extensively. Focal teacher admits less expertise in math. Teacher: "ASAP tests probability, which isn't in district scope and sequence, and we never took a workshop in it."
OPPORTUNITIES FOR RELEVANT PROFESSIONAL DEVELOPMENT	Little course taking, except for teacher pursuing administrative certification. Lack of resources has severely constrained ANY development activities for teachers. Five years earlier ADE staff presented a series of training sessions for teachers in holistic teaching and alternative assessments. But teachers hired since then have not had similar activities, nor has district provided any follow through on earlier training.	Though resources are available to develop teachers and curriculum, they are directed toward activities more consistent with school's basic skills orientation. Teachers say they have not been provided with materials and skills that would allow them to change toward ASAP.	Resources for professional development are ample (1 day /month/teacher), but are directed toward activities consistent with district philosophy (e.g., TQM) rather than ASAP. No professional development in ASAP-consistent instruction was made available to teachers. Staff development is monologues of experts directed toward teachers rather than clinical models. No time for teachers to experiment or collaborate over activities outside the district demands.	Resources for staff development are ample and marshaled toward goals consistent with ASAP. More so for writing and literacy than for math. Absence of competing demands that characterize other sites means that more time is available to experiment, collaborate, and reflect. Teachers serve as resources for each other. Focal teacher has taken a course every semester to enhance her capacity in literacy, literature, integrated curriculum, cooperative learning, etc.

	VALOR	FRANKLIN	PINES	HILLDALE
PRESENCE OF ASAP GATE-KEEPER	ASAP coordinator resigned before end of year. Coordination and communication lacking between anyone knowledgeable and the teachers.	Teachers were unaware there was even a coordinator present in the district. Coordination and communication were thus lacking. No one person took responsibility for informing teachers about ASAP or constructivist education.	ASAP coordinator was school psychologist who reported to the director of testing, a strong CRT and skills-teaching advocate. She provided information about ASAP requirements to teachers, but was not expert in curriculum or alternative assessments. No one teacher or principal stepped forward to be a spokesperson for ASAP. Constructivists were not in leadership positions.	Focal teacher serves as effective gatekeeper. Since the outset of ASAP planning, she has been on state-level and district committees. She regularly attended the state training sessions on ASAP administration and scoring and kept fellow teachers apprised of what the ASAP program meant. She is also trained to be an official scorer.
TEACHER IMAGE OF PUPIL	Concrete-sequential teachers' image of pupils were consistent with district's image: Students are empty containers into which curricular skills are poured. Pupils have innate ability and motivation, irrespective of teachers' efforts. "There is a maturational rate in each individual that they have no control over. And we'll never change that individual who can't grasp the concepts." Learning is a solitary process that occurs within the learner. Constructivist teachers believe students come to school with knowledge, experience, interests that interact with materials and teaching acts.	Teachers' view is consistent with district's: Degree of disadvantage is central. "These pupils" arrive at school with nothing. "Every point they get on those tests, they get from here." Pupils can't transfer or generalize.	Teachers would prefer to teach in constructivist ways, because that is the way pupils learn best (intentional, process- and real-world oriented, interactive, higher order problem-solving, etc.), but these views are suppressed in the dominant district culture, which views pupils in the opposite way. Because of the dissonance, teachers come to accept district image of pupils, that pupil ability is an enduring trait that must be diagnosed and learning activities geared toward. Pupils come to school empty, learn by practicing hierarchi-cally-arranged sub skills. Pupil failure is attributed to personal or family deficiencies.	Pupils' interests and prior knowledge vary; therefore starting instruction from "where kids are" means that appropriate instruction must be constructed by teachers within the context, not standardized or determined from "higher up." Pupils' skills develop unevenly and as a consequence of understanding, not as precursor. Pupils' social nature means that learning should be collaborative. Pupils make meaning together. Learning is active. Prior knowledge and experiences and interests are vital elements in arranging learning opportunities for them.

	VALOR	FRANKLIN	PINES	HILLDALE
TEACHER IMAGE OF TEACHING	Concrete-sequential teachers believe that teachers transmit curriculum as received from external authority. They must maintain the proper conditions of order so that private learning can occur. They must provide opportunities for pupils to practice skills repetitively until mastery is achieved. The teacher "needs to be really focusing on the deficits." Constructivist teachers believe that teachers should design instructional activities that interact with pupil's interests and prior knowledge, and social nature, at least in regard to language, literature, and reading (but not in regard to math).	Irrespective of CRTs, teache[illegible] feel they can "still shut the door" and determine what happens in class (yet they still comply with district requirements). Teacher: "If there were no guidelines for teachers, it would be chaos."	Although teachers maintain some semblance of control over what happens in the classroom, that diminishes under pressure to raise scores by practicing subskills. Teachers would like to teach in ways consistent with ASAP or constructivist models. Some even come into the district with such competencies, but lose them in the shuffle of packed curriculum and test demands. Teacher: "If I'm thorough teaching the basics, I don't have time for this or that." To do more integration, teachers must teach and test faster and more efficiently, otherwise pupils get cheated and basics will get neglected.	Teachers should start from where kids are and the problems or themes pupils find compelling. Teachers should serve as resources in organizing and accessing information sources, encouraging exploration and problem solving, showing the connections between subject matter, and assessing pupils' growth. Teachers need to collaborate with each other and model that process for pupils. Teachers need to look after pupils' social and psychological development and model effective citizenship. Teachers coach, encourage, make good materials available, and provide quality experiences. Teachers don't see themselves as filling deficits, but neither do they ignore problem areas. Teachers need to be self-directed. "It's actually research and data that has changed my teaching techniques" rather than external pressure or presence of ASAP.

	VALOR	FRANKLIN	PINES	HILLDALE
TEACHER IMAGE OF CURRICULUM	Concrete-sequential teachers believe that curriculum is a hierarchically-arranged collection of skills to be mastered. Even supposedly integrated curriculum is "dis-integrated." The goal of learning should be "getting it right." Integrated, holistic science and social studies programs were systematically ignored or "dis-integrated." Constructivist teachers' goals for pupils are to "get it," that is, to understand conceptually. The nature of learning is social and interactive. Students are meaning-makers. When teachers react to pupils' struggles (under conditions of inadequate prior knowledge) with activities that "add sense," learning is meaningful, integrated, and whole.	Curriculum presents each skill sequentially. Repetition is key to mastery. Emphasis on accuracy, not understanding. Anything not on the CRT is regarded as frill, and there is no time for frills when you have to repeat drill on skills. The few teachers who favor whole language feel frustrated and disenfranchised and conflicted. "I don't have to get a reading book to do reading. I can get a science or social studies lesson and that to me is reading for the day. But that's the way I do it, and I know we're not supposed to do it that way."	Teachers feel that there ought to be more to curriculum and teaching than what the district requires or that complying with test mandates demands of them. Teacher: "There's much more to teaching ... how to write and explore your creativity." But diverging from standard curriculum exacts a price. When the formative test results show their students are falling behind the sequence, their anxiety rises and the pace of activity quickens, until they give up and go back to "playing the game," even though that produces "robots." Curriculum is packed, it is something that comes from above and must be covered. "They're giving me more stuff to do, more shit, but they're not giving me any more time to prepare." Nominally integrated units are dis-integrated. Though using a conceptual math program, teacher invented worksheets to drill pupils to mastery on math facts.	With encouragement of principal, teachers endorse a constructivist image of curriculum: whole language and conceptual math. Subject matter should be integrated, address real-world concerns and issues and pupil interests. Much value placed on interactional and cooperative learning opportunities. Teachers use literature-based reading program and Math Their Way and Math a Way of Thinking rather than basals and workbooks. Students form mathematical hypotheses and understanding based on "experiments." Principal tells parents who are "shopping" for schools not to expect seeing many worksheets, as teachers are more interested in authentic problem-solving than drilling to ensure getting the right answer. Writing process permeates and unites all curricular areas. Skills come as a consequence of understanding. Extended project work and thematic units cross disciplinary boundaries and respond to pupils' real-world concerns and interests. No textbooks evident.

	VALOR	FRANKLIN	PINES	HILLDALE
TEACHER IMAGE OF ASSESSMENT	Both concrete-sequential and constructivist teachers distinguish between assessment and testing. For concrete-sequential teachers, assessment amounts to pupils' "displaying" correct answers so that instruction can proceed along the hierarchy of skills; or if incorrect, teachers can regroup pupils and repeat instruction. For constructivist teachers, assessment is teacher judgment of pupil understanding and integration into knowledge structure (particular to each pupil). Both groups define testing as that which is done for external audiences. Regardless of the form of the measure (CRT, ITBS, or performance test), teachers orient toward the accountability function the test serves. Testing is a silent, individualistic "display" of correct responses. ASAP, which is supposed to be integral with instruction, is a separate activity, stressful for both teachers and pupils, an add-on to what they would be doing otherwise, noninteractional and nonteacher mediated.	"Our principal has not supported whole language and doesn't care about ASAP results." ITBS is linguistically and culturally biased. Valid assessment is demonstrated growth, but lack of growth or low scores on tests are due to lack of vocabulary, motivation to do well, not necessarily bad teaching. Teachers know pupils best and therefore should determine what gets assessed. But some think that the CRTs "don't mean a thing."	CRT testing is done for accountability purposes, to satisfy external audiences. Teacher: "Parents said, 'we want better schools, dang it.' [so message from principal] is we need to appease this community," whether that is the right thing to do or not. To teachers, CRT is just a reality of this district, something they have to live with. Principal provides frequent reminders of its importance through comparative graphs, updates, and suggestions for instructional strategies for pulling up scores. Teacher on effects of testing: "We zero in on those ... that's what our district wants us to teach." Teacher: "The first year I didn't know a lot about [CRTs] and I didn't really care. I was teaching very whole language and math manipulatives-based. When the kids' test scores came back, they didn't do very well at all, but I knew we had learned but the test didn't show it. So I added more subskills attention and went with that."	Teachers view testing as fundamentally different from assessment. Testing should be objective to be valid; assessment is valid when it is local, teacher-determined, and shows growth. Role of testing is high-stakes accountability to satisfy external audiences; role of assessment is to advance instruction. Testing is getting the right answer Assessment is more oriented to process, open-ended, not standard responses across pupils and socially constructed. ASAP is something closer to the way teachers teach, so it is more successful and less stressful. ASAP can be a lever to change behaviorist teachers to holistic teachers. There is discontinuity between Forms A, B, C and monitored Form D. A, B, and C are interactive, relaxed, process-oriented, teacher- and peer-mediated, like instruction. D is silent, individual, independent, teacher-centered, disconnected from instruction.

BEST COPY AVAILABLE

	VALOR	FRANKLIN	PINES	HILLDALE
TEACHER IMAGE OF ASSESSMENT (CONTINUED)	Teachers would prefer to have tests play a diagnostic function rather than an accountability function, to help them better understand pupils' learning problems.		ASAP is a step in the right direction, and more in keeping with better instruction, but is still perceived negatively because (a) it is not aligned with district objectives and CRT's, (b) it is administered in spring and reported too late to provide useful information to teachers, (c) comes from ADE as just another externally mandated accountability device, and (d) administrative and scoring problems. Instead, testing needs to be teacher-centered, objective, standardized in administration and scoring. "ASAP is just another thing we have to do."	In Form D, teacher worries about what is to be scored and points pupils' attention to that. Warns them not to spend so much time on drawing (unscored) that they won't have enough time for their story (scored). Teachers believed that ASAP was the state's preferable alternative to ITBS. If ASAP fails, they fear a return to ITBS testing. They feel vindicated that what they perceive to be ASAP's philosophy matches what they perceive to be their own philosophy.

	VALOR	FRANKLIN	PINES	HILLDALE
TEACHER ACCOMMO-DATION / RESISTANCE TO ASAP	Teachers tend to comply with whatever authority dictates.	Teachers tend to comply with whatever authority dictates.	Even teachers who favor holistic instruction come to comply with the district's concrete-sequential model. Teacher: "I've done it. I've done my tap dance." Teacher with personal philosophy most consistent with ASAP resigned in frustration and transferred to an alternative school.	The form of accommodation taken is deliberate and tactical. Teachers and principal see the ASAP program as consistent with the direction they were going independent of the reform. They see it as an inducement to concrete-sequential teachers and top-down administrators to think in different (constructivist) ways about pupils and teaching. On the other hand they are wary of the accountability possibilities inherent (later realized) behind ASAP and attend to that which is to be scored as a possible high-stakes instrument on which they will be expected to score high.

	VALOR	FRANKLIN	PINES	HILLDALE
TEACHER FAMILIARITY WITH ASAP, RUBRICS, ESSENTIAL SKILLS	Only one teacher, who resigned before year's end, was a trained scorer. Others had one on-site ADE workshop. Most had heard of AZ Essential Skills but had not seen or used them except for those lesson plans they had had to code according to sub skills tested on CRTs. Limited use of Forms A, B, and C.	Teachers speak knowledgeably about Essential Skills, because the district CRTs are in part built around them. Only one third-grade teacher seemed to have a strong grasp of ASAP and scoring rubrics.	Teachers received some training during previous year in conjunction with ASAP pilot, but level of awareness is still low. They express understanding of Essential Skills in that district objectives are perceived to have been aligned with them. No one was expert in rubric scoring during this first year. They first got exposure to rubrics during the training they received two weeks before ASAP was to be administered.	Key teachers are well versed in ASAP and scoring rubrics. They thus are savvy about focusing pupils' attention on what is to be scored and have adapted a test-wiseness strategy to maximize scores. Teacher reminded pupils to use complete sentences, for example, because she knew that would make a difference between getting a score of 2 and 3. "Giving the test is much easier if you start with the scoring rubric and work backwards" because that means you are giving the scorers exactly what they want.